Construction of Teaching Performance Evaluation System for College Teachers Based on Data Mining

Hanqiao Sun

Department of Economics and Trade, Shandong Management University, Ding Xiang Road, Jinan, China
Jmxyshq@163.com

Abstract
Data mining has become an important means to improve the level of educational management under the background of big data. In this paper, the CART decision tree algorithm of data mining is applied to the performance management of college teachers. On the basis of statistical data, the factors affecting teaching quality such as teaching attitude, teaching methods and teaching preparation are taken as decision attributes to construct a decision tree. The specific construction process of CART decision tree algorithm in teaching performance evaluation of university teachers is illustrated by examples, and the classification model of teaching performance based on CART tree is generated. It is found that teachers who are fully prepared for teaching often show good teaching attitudes and better teaching performance. For the staff with good teaching quality but low teaching preparation score, if they have good teaching means or good teaching attitude, their teaching performance will be higher; If the teaching method is not skilled or the attitude is not good, the performance is average. It can be seen that teaching performance and teaching quality, teaching preparation, teaching means and teaching attitude are mutually restricting and influencing each other. Finally, targeted and reasonable performance evaluation suggestions are put forward.

Keywords: *data mining, CART decision tree, performance management, college teacher*

1 INTRODUCTION

There are multiple objectives in the evaluation of teachers' performance in colleges and universities. How to measure teachers' performance objectively and accurately is also a difficult and complicated task. At present, economic and institutional incentives are widely used in colleges and universities, and the incentive effect is not significant. By studying the influencing factors of teaching performance, this paper uses decision tree CART algorithm to explore the potential decision rules of teaching performance of university teachers. It tries to find out the key factors that affect their performance and provide effective incentive measures in line with the characteristics of university teachers, so as to improve the overall organizational performance of universities.

From the 1970s, scholars began to study the teacher performance evaluation system, and now the system is gradually developing towards the direction of information, many researchers evaluate its value from the perspective of information system itself. In order to improve the efficiency of the information system and promote the improvement of the operation efficiency, the early research focused on the evaluation of the software response time, stability and efficiency of the information system. Scholars have studied the informatization of teaching management, established models to analyze the informatization management, and predicted the important role played by the informatization of teaching management in promoting teaching evaluation in colleges and universities [1].

In recent years, the teacher management in colleges has made great progress, and some colleges and universities will be able to complete the design and development of personnel management systems by relying on internal technical personnel [4]. Some researchers have put forward their views on the structure of the performance appraisal system. They believe that the performance appraisal evaluation includes the teacher's morality, the mutual cooperation between teachers, the quality of teaching, the interaction between teachers and students, and the risks of teachers' positions. They also introduced some new methods in teacher performance appraisal, such as analytic hierarchy process/fuzzy evaluation method, etc., which provided a method for the objective and scientific evaluation of

B. Fox et al. (Eds.): IC-ICAIE 2022, AHCS 9, pp. 227-232, 2023.
https://doi.org/10.2991/978-94-6463-040-4_34

performance appraisal [5]. However, as many colleges and universities in China are public institutions, there are overlapping functions in the management of various departments. It improves the personnel management system research and development of the difficulty. At the same time, influenced by the teacher personnel management policy, it has a high frequency of adjustment. In addition, the technological strength of colleges and universities and the development of campus information construction are not balanced. Therefore, there are still many deficiencies in teacher performance evaluation.

2 PRESENT SITUATION OF TEACHING PERFORMANCE EVALUATION OF COLLEGE TEACHERS

At present, the teaching performance evaluation of college teachers mainly uses quantifiable institutional evaluation indicators. There are defects in evaluation efficiency, mainly reflected in the following points:

In order to match the requirements of society, industry and post on students' practical working ability, college teachers should accurately grasp the timeliness and practicability of teaching contents in the process of teaching. They need to introduce cutting-edge knowledge, emerging technology and new skills into the curriculum, and constantly enrich the teaching content in the process of classroom teaching and extracurricular training, so as to improve the quality and quality of talents. However, the impact of these teaching preparations on teacher performance is not clear at present.

At the same time, in order to cope with the development of educational science and the application of advanced educational technology, college teachers also need to constantly innovate, introduce new teaching methods, and enrich curriculum teaching with cutting-edge professional technology and knowledge. These teaching methods can promote classroom quality and teaching effect. However, due to the knowledge difference between majors, there is no good evaluation standard for teaching methods.

In addition, college teachers not only play an important role in imparting knowledge, but also play a key role in shaping students' personality and improving students' psychology. College teachers need to help students establish ideals and beliefs by virtue of their own moral cultivation and social experience. Therefore, their own behavior and inner thoughts are the key factors for the success of education. These teaching attitudes are not easy to quantify, so it is difficult to effectively evaluate their impact on performance.

Because the performance management in colleges and universities has not fully realized informatization, it cannot provide effective decision-making basis for the management system of teachers in colleges and universities. There are still many problems related to performance management, which need to be studied. For example, how to evaluate teachers through a set of scientific and reasonable evaluation system and appropriate index weight? How to form an effective teacher incentive mechanism? In order to better build an evaluation system and effectively measure various polymorphic and disorderly factors affecting teacher performance, it is necessary to study teacher teaching performance under big data, which will effectively improve the level of education management and enhance data culture.

3 APPLICATION OF DATA MINING IN PERFORMANCE MANAGEMENT OF COLLEGE TEACHERS

Data mining is a process of extracting hidden, unknown but potentially useful information and knowledge from a large number of incomplete, noisy, fuzzy and random actual application data [3]. Through statistics, analysis, synthesis and reasoning, it summarizes the patterns and internal laws of relevant contents. It helps decision-makers analyze historical data and current data, and find hidden relationships and patterns, so as to predict future behaviors and provide effective support for decision making [6].

This paper mainly applies the decision tree algorithm in data mining technology to teacher performance appraisal, finds out the useful information hidden behind the data, and finds out the potential factors affecting teacher performance appraisal, so as to provide information for relevant functional departments, promote teachers' teaching work better, and improve the quality of teaching and education.

3.1 Decision Tree Algorithm

There are many data mining techniques, including neural network, genetic algorithm, decision tree algorithm, statistical analysis method, rough set method and so on. This paper is mainly based on decision tree, because decision tree is widely used in the field of data mining, especially in the classification problem is a very effective method, and decision tree also has some other advantages [7].

First of all, decision trees can be graphically generated. Decision tree models can be represented by graphs or rules. Graphical analysis results are easy for users to understand, and rules are easy to explain and understand. Since the end users of the data mining system are teaching managers who do not understand the knowledge of data mining, graphical analysis results and rule display make them easy to use. Secondly, decision tree method can deal with continuous or discrete variables. There are many data types involved in data mining system, including not only discrete data but also

continuous data. Among them, there are many discrete data, and the decision tree method is better in dealing with discrete data, and can also deal with continuous data. Thirdly, the classification speed of decision tree method is fast, especially suitable for large-scale data processing. In the decision tree method, the size of the tree is independent of the size of the database, so the calculation of the decision tree method is relatively small compared with other methods. The decision tree can still be constructed when there are many variables in the model. Decision tree method can improve processing speed and execution efficiency of data mining system. Finally, the decision tree method can clearly show the importance of features. The higher the level of the node in the decision tree, the more important the feature represented by the node is, and the less important the feature represented by the node is to the classification. Therefore, the functions of nodes at the same level are basically the same without obvious differences in size. Decision tree method is predictive, so when analyzing a new data object using decision tree, it can infer the corresponding target variable value according to the value of the input variable of the data.

To sum up, according to the advantages of decision tree and the characteristics of data in data mining system, this paper adopts decision tree method to mine the data of performance appraisal, find out the factors affecting teachers' teaching performance, and provide basis for schools to formulate relevant policies.

3.2 Application of Decision Tree in Performance Analysis

The establishment of decision tree model mainly uses a concept of Gini index. The selection strategy for Gini Index is to select the smallest attribute in the Gini Index value as the split node of the decision tree. The CART algorithm based on Gini Index is suitable for less categories Training sets, and the size of generated subsets are similar. The Gini value is used to measure disorder, or the degree of disorder, in a system. The higher the Gini value is, the higher the degree of system chaos is. The purpose of establishing the decision tree model is to reduce the degree of system chaos through appropriate classification, and its calculation formula is as follows:

$$Gini\ (T) = 1 - \sum p_i^2 \tag{1}$$

Where, p_i is the frequency of category appearing in sample T, that is, the ratio of samples of category to the total number of samples.

Σ is the summation formula, the sum of all p_i^2's.

Formula (1) represents all the records in the subset. After the calculation, the Gini value before and after the division of data set S is obtained. The degree of disorder in a system is indicated by the Gini value. The higher the Gini value is, the more disorderly the system is. The lower the Gini value is, the more orderly and clear the system is. Therefore, the development order of teaching performance and decision-making can be clearly found according to the numerical situation.

When a variable used for classification is introduced, the Gini index formula after segmentation is:

$$Gini(T) = \frac{S_1}{S_1 + S_2} Gini(T_1) + \frac{S_2}{S_1 + S_2} Gini(T_2) \tag{2}$$

S_1 and S_2 are the sample sizes of the two categories, and *Gini (T1)* and *Gini (T2)* are the Gini value of the two categories.

The generation strategy of CART decision tree based on Gini Index is as follows:

Firstly, the CART classification method based on Gini Index is used to process the data in the training sample set, and the sub-data sets in the training set are classified first. Secondly, when the class distribution of a certain sub-data set in the training set roughly belongs to the same class, the class represented by this sub-data set can be regarded as a node of the decision tree to stop the continuous expansion of the node and turn the node into a leaf node. Then repeat the process until you end up with a simple and easily understood decision tree.

According to the above calculation, the Gini value is used to divide the decision tree and build the decision tree analysis model as shown in Figure 1.

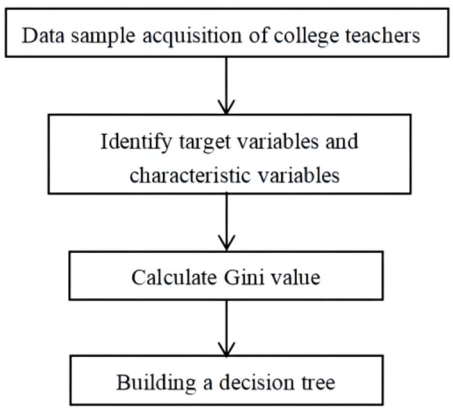

Figure 1: The schematic diagram of the decision tree analysis model

When constructing decision tree analysis model, it is necessary to acquire and sort out the sample data of college teachers firstly. Then the Gini value of target variables and characteristic variables affecting performance are calculated. Finally, the analysis results will be output, so as to complete the decision-making analysis of teaching performance of university teachers.

3.3 Analysis Procedure

In this paper, a total of 513 paper and online questionnaires were issued by taking university teachers as samples, taking advantage of opportunities such as field research, academic conferences and competitions, as well as online platforms such as network platform and mobile phone software. 457 questionnaires were collected, with a recovery rate of 89.08%. After screening and filling invalid questionnaires with missing items, severe extreme reactions or obvious contradictions, 420 valid questionnaires were finally collected, with an effective rate of 91.90%. Gender, age, teaching age, educational background, professional title and other characteristics of personnel were taken into account in the distribution of samples. According to the sample results, the gender ratio is 48.81% male, 51.19% female, and the gender distribution is even. The age structure presents a normal distribution, with teachers with 11-20 years of teaching accounting for the most, nearly 30%, followed by those with less than 10 years of teaching accounting for 23.1%, and those with less than 3 years of teaching accounting for 22.62%. In the educational structure, master's degree accounts for the most, accounting for nearly half of the total, followed by doctoral degree, accounting for 37.86%, and other degrees, accounting for 12.62%. The largest proportion of professional titles is lecturer (37.86%), followed by associate professor (35%), teaching assistant and professor account for the same proportion, and the structure of professional titles is reasonable. It can be seen that the sample size selected in this paper is reasonable, the sample distribution is uniform, the effective sample rate is high, and there is no over-concentration of sample features. The sample composition is shown in Table 1.

Table 1: Respondent profile

Demographic profile		Percent (%) (N = 420)
Gender	Male	48.81%
	Female	51.19%
Marriage	Married	88.57%
	Unmarried	11.43%
Age	Under 30 years	11.43%
	31-35 years	21.43%
	36-45 years	37.14%
	46-55 years	20.95%
	55 years or over	9.05%
Teaching years	Under 3 years	22.62%
	4-10 years	23.10%
	11-20 years	29.76%
	21-30 years	15.24%
	31 years or over	9.29%
Degree	Bachelor	12.62%
	Master	49.52%
	Doctor	37.86%
Title	Assistant	13.81%
	Lectorate	37.86%
	Associate professor	35.00%
	Professor	13.33%

To ensure the validity and reliability of the scale, this study combined the existing mature scale and conducted in-depth interviews with 7 human resource managers in colleges and universities and 15 front-line teachers, 2 experts in human resource, and 1 professor of sociology from 7 universities in China. The relevant variables were collated and mined, and a questionnaire was designed, containing 24 questions. After the initial test of the small sample, a scale consisting of 20 items was finally formed according to the results combined with the suggestions of expert discussions.

Likert five-point scale was used to classify the options into "completely inconsistent", "inconsistent", "uncertain", "consistent" and "completely consistent", and give 1, 2, 3, 4 and 5 points respectively. Of teaching attitude, teaching means, teaching preparation, respectively set up scale item [2].

According to the weight of the statistical score, the project score comprehensively weighted to the teaching attitude, teaching methods, teaching means of three grades. Teaching attitude is divided into three levels: excellent, medium and poor, and teaching methods are divided into three levels: excellent, medium and poor. Teaching preparation is divided into two grades: excellent and poor. The teaching performance test table constructed from sample data is as follows:

Table 2: Teaching performance examination form for college teachers

Teacher's No.	Attitude	Methods	Prepara-tion	Quality
01	medium	medium	good	qualified

02	medium	poor	good	excellent
03	medium	poor	poor	qualified
04	poor	good	good	qualified
05	good	good	poor	excellent
06	medium	good	good	qualified
07	good	poor	good	qualified
08	medium	medium	poor	qualified
09	poor	medium	good	qualified
10	good	good	poor	excellent
…	…	…	…	…

Among the 420 teachers, 186 were rated as excellent, while the rest were qualified. The Gini value was 0.494, indicating a high degree of confusion.

Gini values of teaching attitude and teaching means on teaching quality were calculated respectively:

$$\text{Gini}(D, \text{attitude} = \text{excellent})$$

$$= \frac{120}{420}\left(1 - \left(\frac{52}{120}\right)^2 - \left(\frac{68}{120}\right)^2\right) + \frac{300}{420}\left(1 - \left(\frac{134}{300}\right)^2 - \left(\frac{116}{300}\right)^2\right) = 0.493$$

Gini(D, attitude = medium) = 0.493

Gini(D, attitude = poor) = 0.491

Gini(D, methods = excellent) = 0.455

Gini(D, methods = medium) = 0.475

Gini(D, methods = poor) = 0.492

Gini(D, preparation) = 0.49

The one with the smallest Gini value is the teaching method, which is used as the root node to expand downward. Similarly, the decision tree is finally generated as follows:

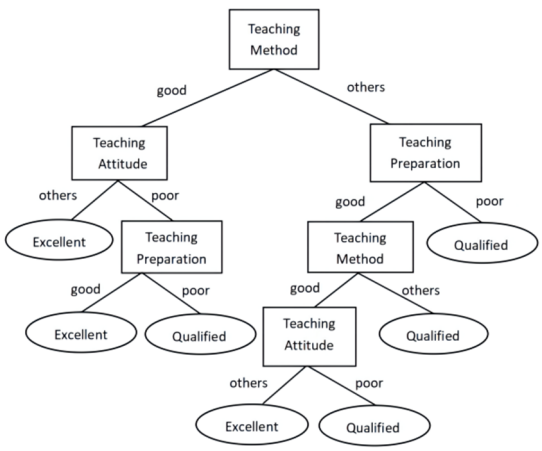

Figure 2: The final decision tree that is generated

4 CONCLUSIONS

Through the analysis, the law of teaching performance of teachers in colleges and universities can be obtained: teaching preparation is positively correlated with teaching performance, and teachers with high teaching preparation score tend to show good teaching attitude. Teachers with inadequate teaching preparation, such as low teaching quality scores, will definitely have poor performance. For the staff with good teaching quality but low teaching preparation score, if they have good teaching means or good teaching attitude, their teaching performance will be higher; If the teaching method is not skilled or the attitude is not good, the performance is average. The situation of teachers with poor teaching quality and better teaching preparation is complicated. If they have good teaching methods and teaching attitudes, the heavy teaching tasks may lead to poor teaching quality, so they are still likely to obtain better performance evaluation. Thus, teaching performance and teaching quality, teaching preparation, teaching means and teaching attitude are mutually restricting and influencing each other.

Finally, the application of data mining technology in the performance management of university teachers plays the following important roles in the improvement of teaching level:

1. Find out the real reasons affecting teachers' performance, and formulate corresponding measures and incentive policies based on them, so as to make them more engaged in their work, encourage teachers to develop their own potential and improve their job satisfaction, so as to continuously improve their work efficiency and ability level and enhance their cohesion.

2. Discover performance rules, improve the effectiveness of performance management, support management decisions, guide colleges and universities to evaluate teachers' performance, and help decision-makers in talent planning.

3. Find and implement effective ways of organizational performance management in colleges and universities, improve performance management strategies, help colleges and universities achieve sustainable development of their performance, promote the formation of a scientific evaluation system, improve organizational performance, and improve the level of education and teaching.

ACKNOWLEDGEMENTS

Supported by a project grant from Shandong Management University Research Enlightenment Program "Research on Performance Evaluation of Teachers in Applied Universities Based on Intrinsic Motivation" (Grant No. QH2021R09).

Supported by Ministry of Education "Industry-College Cooperation and Collaborative Education" project named "Construction of practical training course system for investment specialty for application-oriented talents cultivation" (Grant No. 202002220002).

REFERENCES

[1] Alexander Kroll. (2015) Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research. *Public Performance & Management Review*. 3,118-126.

[2] Mark A. Griffin, Andrew Neal, Sharon K. Parker. (2007) A new model of work role performance: positive behavior in uncertain and interdependent contexts. Academy of Management Journal, 4, 949-969.

[3] Marijn Janssen, VDV Haiko, Agung Wahyudi. (2017). Factors influencing big data decision-making quality, Journal of Business Research.

[4] Saqib Shamim, Jing Zeng, Syed Muhammad Shariq, Zaheer Khan. (2018). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. Information & Management.

[5] Yingjuan Sun. (2011) Research on classification method based on rough set, Jilin University.

[6] T. Manjunath Kumar and R. Murugeswari. (2019). Predicting Faculty Performance in Higher Education using Machine Learning. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(4): 9472-9478.

[7] Zhang Liang, Ning Qian. (2015). Two Improvements and Applications of CART Decision Tree. Computer Engineering and Design, 36(05): 1209-1213.